

July 1, 2007

To: Rutgers Business School Faculty
From: Glenn Shafer

Using GMAT/GRE scores for doctoral admissions

In an earlier memorandum, I summarized data showing that graduates of our Ph.D. program who take the most prestigious teaching positions have lower GMAT and GRE scores, on average, than those who take less prestigious positions.

In this memorandum I respond to questions from several faculty members by giving more detail on the data and its implications. I also make the following recommendations for changes in our current practices:

1. Continue to use GMAT and GRE scores, especially in cases where other information on a candidate for admission is limited, but give these scores less weight relative to other information than we have been doing.
2. Increase the weight we give other evidence of achievement and creativity in English writing and speaking. For example, when there is strong evidence that a candidate has been effective in professional roles that require careful writing in English, consider the candidate seriously even if his or her percentiles fall below 50%.
3. When considering the GMAT or GRE scores of foreign students, give more weight to the verbal percentile and less to the quantitative percentile than we have been doing. The median verbal percentile for our foreign students is now about 80%, while the median quantitative percentile is about 90%. We should try to bring these two numbers together.
4. When we admit a student who is weak in English or writing because of their other strengths, insist that the student take full advantage of opportunities for improving their verbal skills, including the tutoring offered by the Rutgers-Newark Writing Center and opportunities in teacher training.

In my judgment, these changes are important not only for improving our placement but also for enhancing our program's contribution to the faculty's research and to the school's teaching mission.

1. Why look at the correlation between performance and standardized test scores in the population consisting of our students?

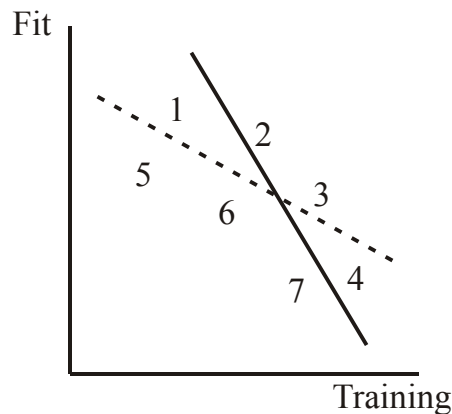
What do the standardized tests measure? According to ETS, the GRE "measures verbal reasoning, quantitative reasoning, and critical thinking and analytical writing skills that have been acquired over a long period of time and that are not related to any specific field of study." It does this imperfectly for US residents and poorly for relative newcomers to the English language and United States culture. But it is still one useful indication of a student's promise, and for this reason we require that all applicants provide either a GRE or a GMAT score.

I made the final decision on admission for most of the students who entered our program after 1998. I want to know whether I used the GMAT/GRE scores as well as possible in making these decisions. Did I give the GMAT/GRE scores the right weight relative to other factors? This is why I am interested in the correlation of the scores with later performance in the population consisting of our students and other applicants similar to them, applicants I might have recruited.

If my analysis had shown that those performing well had about the same GMAT/GRE score on average as those performing poorly, I would have concluded that I had it just right. If I had found that those performing well had higher scores, I would have concluded that I had put too little weight on the scores. Apparently the contrary is true. I have been putting too much weight on the scores.

A simplified example may clarify why I come to this conclusion. Suppose our only goal is to have as many graduates as possible. Suppose we base admissions on only two factors that might affect whether a student graduates: a measurement T of prior training, and a measurement F of the fit between the student's interests and those of the faculty. Perhaps T is a test score, and F is a score I assign after reading the applicant's personal statement. Suppose that when one of the factors is held constant, the chance of graduation increases as the other factor increases.

Suppose we have seven applicants, with the values of (T,F) shown in the diagram below. We are allowed to admit only three. I decide to admit applicants 2, 3, and 4, so that the solid line divides the admitted from the not admitted. Suppose 2 and 3 graduate, but 4 drops out, having decided that he is not interested in our faculty's research after all. What do we think now? We think that perhaps we should have admitted 1 instead of 4, so that the dotted line would have separated the admitted from the not admitted. We should have weighted prior training T less and fit F more.



We can make this picture into a mathematical theory if we want. Under reasonable assumptions, with many applicants and a substantial number admitted, the theory will conclude that if dropouts have higher values of T than graduates on average, then we would have had more graduates had we admitted fewer students like those who dropped out (relatively high T and low F) and more like those who graduated (relatively low T and high F). This conclusion cannot be refuted by speculation about factors other than T and F that might affect the dropout rate, because we select the students from the applicants ourselves, using only T and F.

2. Results I reported earlier (with a typo corrected)

As I reported earlier, GMAT/GRE scores for students we place in universities classified as “national universities” by US News are lower than the scores for those we place in regional and local universities. So if placement in top universities is our goal, we give too much weight to GMAT/GRE scores relative to other factors we consider.

A total of 115 students entered our doctoral program from Fall 1998 through Fall 2003. Of these, 32 dropped out, while 83 graduated or are expected to do so in the next year. I will call the latter group “graduates.” Our staff found GMAT or GRE scores in our archives for 107 of the 115 students. This included 28 of the 32 dropouts and 79 of the 83 graduates.

To get a single score for each student, I averaged the verbal and quantitative percentiles. If the student reported more than one exam result, I used the highest verbal percentile and the highest quantitative percentile.

Because the resulting student scores are skewed, I calculated medians rather than averages of these scores. The median score for all 107 students was 80.75. Table 1 shows additional medians: for the 79 graduates, the 28 dropouts, the 50 graduates who found tenure-track jobs, the 29 who did not, etc.

The median student in our program is in about the 80th percentile on the standardized test they took. This is the median percentile for graduates as well as for dropouts. It is also the median percentile for those who take tenure-track jobs, and for those who take tenure-track jobs in the United States. But the median for those who land jobs in top universities is about 5 points lower.

	#	(v+q)/2
Graduates	79	81
Dropouts	28	80.5
Total	107	80.75

Of the 79 graduates,

with tenure-track jobs	50	80.75
without tenure-track jobs	29	82.5

Of the 50 with tenure-track jobs,

in US	44	81.75
-------	-----------	-------

Of the 44 with US tenure-track jobs,

in national universities	13	75.5
--------------------------	-----------	------

Table 1. Median scores for students in various categories. The number of students in each category is shown in bold. I calculated a score for each student by averaging the student’s verbal and quantitative percentiles, and I report the medians of these scores in the column on the right.

3. Why was my judgment faulty?

Why did I put too much weight on GMAT/GRE scores? Why didn't I put more weight on grades, recommendations, record of accomplishment, commitment to scholarship, and fit with the research interests of our faculty?

Perhaps I weigh GMAT/GRE scores too heavily because they are the only indicators I can compare easily across our many diverse applicants. It is difficult to compare grades in different courses in different programs in different universities. I often know little about the programs or even the universities, even those in the United States. It can be difficult to discern the meaning of a letter of recommendation or the relevance of a prior accomplishment. The GMAT/GRE score seems more objective and hence more trustworthy than my reactions to this other information.

I have personally benefited from standardized tests. Because of my SAT scores, I entered an elite university after graduating from a rural high school. I think I take test scores with a large grain of salt. But is this really true? Perhaps I cannot help but see myself in a student who scores well on standardized tests. In the spirit of full disclosure, I should also acknowledge that I am part of the system that produces these tests. For over thirty years, I have interacted with fellow researchers at the Educational Testing Service, which produces the SAT and GRE tests. I have accepted honoraria for speaking at ETS and serving on the advisory board of their research division. Like other business-school admissions officers, I have also attended GMAC conferences underwritten by the fees students pay to take the GMAT test (I enjoyed Disneyland).

I am also influenced, quite properly, by my colleagues. I usually admit students only after hearing from doctoral coordinators and departmental admission committees. Perhaps some of these colleagues have some of the same biases I have.

Finally, there are incessant institutional pressures to raise the GMAT/GRE scores of our entering students. Several times a year, our program office is asked to calculate average entering scores, which are then passed on to accrediting agencies, to profit-making enterprises such as US News, or simply up the chain of command in the university. These requests feel to us like messages that we should raise the scores if we want the university to think we are doing our job well. Often the message is explicit. Two years ago, an external evaluation of our school complained that the average GMAT of our entering doctoral students is lower than in unnamed competing business schools, and this criticism continues to be echoed in the university.

Provost Steven Diner has pointed out that standardized tests were introduced to give a chance to talented students whose disadvantaged situation limited the other credentials they could garner (see the attached article, which will appear in a book on equity in higher education). Now they play a very different role. They are used to evaluate universities as much as students, and they may actually give an edge to students with the resources to pay for instruction in test-taking.

4. A more detailed analysis

Several colleagues asked additional questions. What happens when we consider only domestic students – say those born in the United States? Are there other factors that predict whether a student will land a job in a national university? What happens when we try to predict performance with the verbal percentile alone or the quantitative percentile alone? Are there differences between departments? Do students who land more prestigious non-academic jobs have higher GMAT/GRE scores? These are good questions, and some of them can be answered.

The tables on the next page give medians for verbal and quantitative percentiles alone, a separate analysis for US born students, and medians for the four large departments.

Is the picture any different for US born students? For US born students, the median GMAT/GRE score for those who drop out is notably higher than for those who complete the program. It is also somewhat higher for those who take tenure-track jobs than for those who do not. Otherwise, the pattern is quite similar to that for the other students.

What other factor predicts whether a student will land a job in a national university? The obvious predictor is whether the student is US born. For US born students, 50% of those taking tenure-track jobs in the US land in national universities (10 out of 20). For the other students, the fraction is 12.5% (3 out of 24).

What do we see when we look separately at the verbal percentile and the quantitative percentile? The median quantitative percentile is about 90, while the median verbal percentile is about 80. For US born students, the balance is reversed, with the quantitative percentile about 75 and the verbal percentile 85. There are two notable deviations from this picture.

1. As we have already noted, US born dropouts have higher scores than US born graduates. This is especially notable for the verbal percentile. When we look at all students, including both foreign and US born, we see the opposite: dropouts have sharply lower verbal percentiles than graduates.
2. The disparity between verbal and quantitative scores disappears for students who land jobs in national universities. For this group, the verbal and quantitative percentiles are both close to 73.

Some of these observations are shaky, however, because they involve small numbers of students. There are only seven US born dropouts, for example.

Are there differences between departments? As Table 3 shows, the departments differ in the proportion of US born students they admit, in their success in placing students in national universities, and in median verbal and quantitative percentiles. Although I do not show the numbers, I did look at the breakdown into graduates vs. dropouts, tenure-track vs. not, and whether the tenure-track job was in the US, and I found the same stability for $(v+q)/2$ over these categories as observed in Table 1 for the program as a whole. I did not look at the verbal and quantitative percentiles separately by department, because there are too few students in each department to give reliable insight into the issues that arise from looking at these percentiles for the program as a whole.

Do students who land more prestigious non-academic jobs have higher GMAT/GRE scores? This question cannot be answered definitively, because we have no good measure of prestige for non-academic jobs, and the number of students involved is very small. Most of the 29 graduates who do not have tenure-track jobs are still on the market or in temporary positions. At most 3 or 4 are in non-university research jobs our faculty would consider prestigious, and all I can say about them is that their GMAT/GRE scores are not uniformly high.

	All students				US born students			
	#	v	q	(v+q)/2	#	v	q	(v+q)/2
Graduates	79	82	92	81	28	82.5	72.5	77
Dropouts	28	71	90	80.5	7	96	84	91.5
Total	107	80	90.5	80.75	35	85	75	79.5

Of the graduates,

with tenure-track jobs	50	82	92	80.75	20	87.5	73.5	79.75
without tenure-track jobs	29	82	92	82.5	8	80.5	70.5	75.75

Of those with tenure-track jobs,

in US	44	82	90	81.75	20	87.5	73.5	79.75
-------	-----------	----	----	-------	-----------	------	------	-------

Of those with US tenure-track jobs,

in national universities	13	72	74	75.5	10	74	72	74
--------------------------	-----------	----	----	------	-----------	----	----	----

Table 2. More medians: all students and US born students, verbal and quantitative GMAT/GRE percentiles. The numbers in bold are numbers of students in the category. The columns headed “v” and “q” give the median verbal and quantitative percentiles, respectively, for these students. The columns headed “(v+q)/2” give the median of the average of the two percentiles, the score used in the initial analysis.

The reader may be puzzled that the median for (v+q)/2 sometimes falls outside the interval between the median for v and that for q. To see how easily this can happen, imagine that Tom, Dick, and Harry have verbal percentiles 10, 50, and 60, respectively, and quantitative percentiles 80, 50, and 20. Then their values of (v+q)/2 are 45, 50, and 40. The medians are then 50 for v, 50 for q, and 45 for (v+q)/2.

Department	v	q	(v+q)/2	Number of students	Number born in US	Number placed in national universities
ABEIS (Accounting)	80	87	78	19	8	2
FE (Finance)	92	97	93.5	24	6	1
MGB (Management)	78	76	75	29	17	7
MSIS (Mgt science)	70	95	80.75	30	4	3

Table 3. Medians for GMAT scores by department. Only the four large departments are included. The three students placed in national universities by the MSIS department were foreign born. The ten students placed in national universities by the other departments were US born.

5. Interpretation

In order to preserve the confidentiality of information about specific students and keep the numbers large enough to make conclusions meaningful, I have given only medians for relatively coarse categories. My interpretation will also use additional information, some of it in the raw data on GMAT/GRE scores, and some of it from my experience administering the program.

Why are both verbal and quantitative GMAT/GRE percentiles lower for students we place in national universities? Consideration of this question must begin with two details in the data.

1. As Table 3 indicates, the majority of students placed in national universities were in the MGB department, and this department tends to admit students with lower percentiles than other departments. It is also true, although this information is not in Table 3, that these phenomena are driven by the department's Organization Management major rather than its International Business major. (The two majors were handled by separate departments when most of these students were admitted.) All the department's national university placements were in Organization Management rather than International Business, and GMAT/GRE scores are rather lower on average for OM than IB.
2. It is also true that the majority of the national placements were US born students, and Organization Management is our only major in which US born students were in the majority.

Interpretation of the results turns on the extent to which the strategy pursued by the Organization Management major when these students were admitted is available to the other majors. Do the other departments and majors have available to them applicants with stronger communication skills (these applicants tend more often to be US born) who can better compete in the national market in spite of having lower verbal and quantitative GMAT/GRE percentiles?

We discard rejected applications after one year. So there is no archival evidence that might get us started on an answer to this question. I can only offer my own opinion. I think the Finance department does receive a fair number of the types of applications needed. Accounting, as Table 3 suggests, falls about halfway between Finance and Organization Management in its willingness to accept students with strong communication skills and weak GMAT/GRE scores and in its success in placing students. The MSIS department, on the other hand, has not had the pool of applicants it would need to shift its strategy, and changes in its doctoral program may be needed to generate such a pool.

Why are GMAT/GRE scores higher for US born dropouts than for US born graduates? We might be tempted to say that students with higher GMAT/GRE scores are more talented and therefore have more alternatives. This formulation relies, however, on the erroneous and misleading assumption that our students are a random sample from the population of everyone who takes the tests.¹ It is more accurate, meaningful, and useful to say that our admissions policies for US born students have sometimes emphasized high GMAT/GRE scores over

¹ People sometimes try to infer something about the whole population of people who take standardized tests from the relation between test scores and subsequent performance for students in particular programs. This is obviously impossible, because the students in the programs are not a random sample from the population taking the tests. Doug Jones alluded to this selection bias in an e-mail in response to my earlier memorandum.

suitability for our program. This has not happened often; there were only seven US born dropouts, and they did not all have high GMAT/GRE scores. But it should make us think twice.

Why are quantitative percentiles higher than verbal percentiles for our foreign students? The US born students we admit tend to have higher verbal than quantitative percentiles, but for foreign students the balance goes the other way. Most notably, foreign born dropouts have a median quantitative percentile nearly 20 points higher than their median verbal percentile.

I suspect that in the case of foreign applicants, we often discount verbal scores relative to quantitative scores, thinking that the verbal skills can be more easily improved when the student arrives in our country. But we do too little to encourage the required improvement. Instead, we throw the student into courses that are relatively quantitative, usually much more so than the actual research work they must do to complete our program. We do offer intensive and effective instruction to improve verbal skills, including our PALS English courses, the tutoring by the Rutgers-Newark Writing Center, and Professor Gardiner's course in Effective College Teaching, but the relatively collegial relationship that most advisers cultivate with their doctoral students makes it difficult for them to insist on the students taking advantage of these opportunities.

6. Context and sensitivities

To avoid misunderstandings, it is important to emphasize that all our students have been admitted because they have excelled on many dimensions. The admissions process must consider many dimensions, because both the skills and attitudes needed for success and the evidence for these skills and attitudes are multidimensional. We go seriously wrong whenever we fall into thinking that applicants are arrayed along a one-dimensional scale of ability. Instead, we should think of them as arrayed in a multi-dimensional space, with almost everyone we consider seriously on the frontier consisting of those who excel on some combination of dimensions, just as applicants 1, 2, 3, and 4 are on the frontier in the diagram we considered in Section 1.

Discussions of the weighting of the different dimensions are nevertheless thorny. In the case of GMAT/GRE scores, the discussion can quickly encounter ethnic sensitivities, because average performance on standardized examinations such as the GRE and GMAT varies sharply across countries. This is well known and widely discussed. In fact, ETS now distributes data on median scores for different countries to admissions officers.

My recommendation that we shift some weight from standardized scores to other evidence of achievement and creativity in English writing and speaking translates, on a practical level, to a recommendation that we should try to recruit more US born students. This is not a new recommendation, and in the past it has led to concern – from our foreign-born students and even sometimes our foreign-born faculty – that their talents and contributions are being disparaged.

I want to emphasize, therefore, that the contribution made by foreign-born students to our program is immense and essential. Like most other US doctoral programs in science, engineering, and business, we must import talent from abroad to survive. For decades, the majority of students entering our doctoral program have come from other countries. It is not realistic or desirable to change this in the foreseeable future. I do think, however, that we should increase the size of the minority of US born students.

To provide context, let me call attention to the balance between US and foreign born students among our current students and among the incoming class.

- 47 continuing students will be in their 2nd, 3rd, or 4th year in 2007-2008. Of these, 14, or about 30%, are US born. If we look at the 33 of the continuing students who are supported by Rutgers (TAships or faculty grants), we find that only 6, or 18%, are US born.
- 25 new students plan to join us in Fall 2007. Of these 10, or 40%, are US born. The others are from China (5), South Korea (2), Egypt, Germany, Indonesia, Iran, Kuwait, Pakistan, Taiwan, and Thailand. If we look at the 14 of the new students who will be supported by Rutgers, we find that 6, or about 43%, are US born.

The greater proportion of US born students in the entering class for Fall 2007 is partly due to a decision on my part to put less emphasis on GMAT/GRE scores and more on other information relevant to success in our program. I recommend continuing this new emphasis.

Although I have not consulted with them on this issue recently, the department chairs, doctoral coordinators, and RBS deans have generally supported over the years the goal of decreasing our emphasis on standardized scores and increasing the proportion of US born students. This goal is sometimes less strongly supported by faculty members who have not had an equally broad experience of how applicants with different characteristics turn out in terms of their contributions to the school and their later placement. These faculty members tend to be more impressed by the apparent objectivity of standardized scores, and they tend to be unaware that the many of our best placements have been students who scored in the lower half of the scale for these tests. The pervasive use of standardized scores to evaluate institutions makes it difficult to share, believe, or remember this information.

We must rely on a large number of faculty members to help evaluate applications. Some are very experienced, some less so. Some serve on departmental admission committees; some provide advice on an ad hoc basis. I hope the information provided here will help them make their contributions more effective.

7. Recommendations

At the beginning of this memorandum, I offered four recommendations.

1. *Continue to use GMAT and GRE scores, especially in cases where other information on a candidate for admission is limited, but give these scores less weight relative to other information than we have been doing.*

This will require changes in our admission process. The program office now initiates deliberation by the departmental admission committees in March by providing electronic copies of the full applications together with an Excel spreadsheet with summary information: degrees earned, grade point averages, GMAT or GRE scores, etc. Recommendations sometimes come back to me with little information about the committee's thinking. To make sure information other than the GMAT/GRE scores is given appropriate weight, we might ask the department to provide a sentence or two for every recommended applicant concerning (1) needed verbal reasoning and communication skills, (2) needed quantitative reasoning skills, (3) accomplishment in and commitment to research, and (4) fit with the faculty.

A further step, which might do much more to strengthen our doctoral student recruiting, would be to constitute the departmental committees in the fall of each year as *recruiting* committees, not merely *admission* committees, charged with contacting masters programs at Rutgers and nearby universities to encourage applications and inviting potential candidates to attend the departmental seminar and meet the faculty. A committee that has already thought about the applicants we want will be in a better position to evaluate applications that appear later in the year and to swing into action when we need to court applicants.

2. *Increase the weight we give other evidence of achievement and creativity in English writing and speaking.*

For every applicant we admit, we need to identify adequate evidence of prior effectiveness in roles that require careful use of English. This will preclude the argument that a person's quantitative accomplishments are evidence of a level of intelligence sufficient to enable them to become a good communicator.

3. *When considering the GMAT or GRE scores of foreign students, give more weight to the verbal percentile and less to the quantitative percentile than we have been doing.*

The balance between verbal and quantitative skills in admissions to doctoral programs is often a matter of contention in business schools and economics departments. The quantitative skills are important for foundational courses everyone agrees are essential, but the verbal skills are ultimately more important, for they are most important in research, teaching, and self-presentation to colleagues and administrators. If those who teach the foundational courses control admissions, there can be a debilitating emphasis on quantitative skills. Each department needs to think about this issue carefully. Those who teach foundational courses may need to rise to the challenge of doing so in a more verbal way. If the MSIS department continues to play a central role in teaching many of the quantitative foundational courses for the other departments, we need to find a way of engaging those who teach these courses in the statistical and modeling aspects of research in the other departments, so that the verbal presentation helps the student make the bridge to dissertation research.

4. *When we admit a student who is weak in English or writing because of their other strengths, insist that the student take full advantage of opportunities for improving their verbal skills, including the tutoring offered by the Rutgers-Newark Writing Center and opportunities in teacher training.*

We may need to provide more explicit authorization to the doctoral coordinators to review each student's performance annually and recommend requisite work.

In a previous memorandum, I argued that the doctoral program's contribution to the faculty's research and the school's teaching mission is more important than placing graduates in prestigious positions. I believe it is worthwhile for the faculty to make the effort these changes will require because the changes will also further substantially these other goals.